Learning from the outcomes of others: Stock market experiences of local peers and new investors' market entry*

Markku Kaustia

Helsinki School of Economics

Samuli Knüpfer

London Business School

First version: February 15, 2008

This version: June 30, 2009

Abstract

We find that an individual's stock market entry decision is positively influenced by the stock market performance of local peers in the previous month. This neighborhood return effect is economically significant, and stronger than the effect of general market returns. The effect is limited to positive outcomes, negative peer returns have no effect. The future return of the stocks bought by the new investors is a declining function of the local peers' past performance. These findings are consistent with naïve extrapolation from the success of peers, and selective communication, i.e., people do not talk about decisions that have produced inferior outcomes.

Keywords: Investor behavior, social interaction, social learning, stock market participation

JEL classification: G11, G24, D83

* We thank James Choi, Lauren Cohen, David Faro, Annette Vissing-Jørgensen, Motohiro Yogo, and the seminar participants at the Helsinki School of Economics and London Business School for comments. Antti Lehtinen provided excellent research assistance. Financial support from the Finnish Foundation for Advancement of Securities Markets is gratefully acknowledged.

1. Introduction

> "That others have made a lot of money appears to many people as the most persuasive ... evidence that outweighs even the most carefully reasoned argument..."
>
> Robert Shiller, Irrational Exuberance, 2005.

Imperfect social learning from outcomes experienced by peers can slow down the diffusion of information.[1] In many circumstances peer outcomes are not directly observable. Learning from outcomes in such circumstances requires communication, which may not take place, or may be biased. Even in cases in which outcomes are directly observable, social learning can be hindered. Unobservable individual characteristics of peers can impact the success of using a new technology: what works for your peer does not necessarily work for you. In many environments outcomes also contain a stochastic component, which makes outcomes noisy signals of decision quality.

Social learning thus involves a chain of events: action, outcome, communication, and interpretation. There is very little empirical research to guide our understanding of that chain of events.[2] In this paper we conduct a large-scale empirical test of outcome-based social learning. Our setting is the stock market. We test whether the returns experienced by the existing investors in a given neighborhood affect the likelihood of new investors to enter the stock market in the same neighborhood. In this setting we are able to define the outcomes experienced by the peer group (stock market performance) explicitly, and measure them without error in the data.

Some features of the setting suggest that little evidence of outcome-based social learning might be found. People cannot directly observe the stock market performance of their neighbors, but have to rely on indirect cues, such as verbal accounts. Communication may not be truthful if appearing to be a successful investor carries private benefits. Even if the true return on a

[1] For theoretical treatments, see Ellison and Fudenberg (1993; 1995), McFadden and Train (1996), Persons and Warther (1997), and Banerjee and Fudenberg (2004).

[2] The studies that we are aware of are in the areas of agricultural and development economics, where the environment is more deterministic. We review the literature in Section 2.

neighbor's portfolio were known, it would still be a noisy indicator of investment skill. It is thus difficult to infer whether the neighbor's investment strategy is worth following.

Surprisingly, we find strong evidence of outcome-based social learning. We use data on the stock holdings and trades of the entire population of individual investors in Finland.[3] We first establish the effect of social interaction in the tendency of new investors to enter the stock market utilizing a natural experiment. An exogenous shock to individuals' share ownership occurred as a local mutually owned telecom operator converted into a public corporation, and its customers received shares in the company.[4] The new shareholders enjoyed a windfall gain as the shares were listed. Difference-in-differences estimation shows that people who live in the same neighborhoods, but who were not entitled to receive these shares, were significantly more likely to enter the stock market shortly after their neighbors' shares were listed in the exchange.

In the main empirical tests we aggregate direct stock holdings at the zip code level and measure the monthly returns of these zip code portfolios. This introduces substantial variation in investors' performance, which was lacking in the natural experiment. We find that high stock market returns during a month, in a neighborhood stock portfolio, are associated with an increase in the number of new investors entering the stock market in the same neighborhood, in the following month. Our data and method allow identifying the neighborhood effect as it relates the outcomes of one group (existing investors) to the actions of another, distinct group (potential new investors). In addition, the dynamic feature eliminates concerns for the effect of common unobservables. We control for month and zip code fixed effects in the analysis, so the results are not driven by general market trends, or time-invariant regional characteristics. The effects are economically large: an increase from the median neighborhood return to the 75th percentile return causes an annual increase of several percentage points in the participation rate. We also directly contrast the neighborhood return effect with the general market return effect, which in itself is a significant predictor of entry. We find that the neighborhood effect is more powerful.

In addition to testing the main hypothesis on the impact of outcome-based learning, we investigate the channels through which outcomes affect social learning. Shiller (1984; 1990) suggests that people naïvely extrapolate from their neighbors' success in the stock market. Under

[3] The data source, the Finnish Central Securities Depository, is an official ownership registry, and the data are thus very accurate and reliable. More description of a subset of the data is provided in Grinblatt and Keloharju (2000).

[4] Kaustia and Torstila (2008) use a similar setting to study the effect of stock ownership on political views.

this hypothesis, people have full information about whether their neighbors are participating in the stock market, and they update their stock market return expectations based on their neighbors' performance. We call this the 'extrapolative expectations' hypothesis. The second hypothesis relaxes the assumption of full information about the peers' stock market participation status. Under this hypothesis, individuals are initially not aware of whether their peers participate in the stock market, but become aware of this when people talk. Learning about the peers' participation can make prospective investors want to enter the stock market, for example due to relative wealth concerns. Due to an incentive to appear competent, or because of self-serving bias in recall and attention, people are likely to talk more about their investments when their performance has been good. This hypothesis is called the 'selective communication' hypothesis. These hypotheses are not mutually exclusive, however. In particular, it is possible that communication is selective, but naïve extrapolation could still occur when the information is transmitted.

We find a strong asymmetric relation between returns and entry: positive returns increase entry, but negative returns do not decrease entry. This pattern is consistent with the selective communication hypothesis. That is, people are more likely to discuss their stock market experiences with others when those experiences have been favorable. Importantly, the pattern is not explained by the naïve extrapolation hypothesis alone. Naïve extrapolation predicts that people update their stock return expectations based on their neighbors' experience. But under naïve extrapolation the neighborhood return effect should be symmetric, and thus, contrary to what we observe, negative neighborhood returns should reduce the willingness to enter the market. Selective communication, on the other hand, predicts this asymmetric pattern whether the underlying mechanism is based on expectation or preference interactions. Selective communication can thus lead investors to behave "as if" they naïvely extrapolate from their peers' success.

Finally, we investigate the performance of the new market entrants. Existing studies find that individual investors make value-reducing trading decisions while institutional investors gain at their expense (Odean, 1999; Barber and Odean, 2000; Barber et al., 2009). New investors who naïvely extrapolate from their peers' success should be particularly prone to making inferior choices. We test this hypothesis by regressing the performance of the new investors' initial portfolio to past neighborhood returns, controlling for fixed time effects. We find that this relation is strongly negative. This implies that investors who enter the market after their neighbors enjoyed good returns tend to buy precisely those stocks which underperform from that point forward. This is consistent with naïve extrapolation.

There are two alternative explanations to our findings that do not involve social learning. Both rely on investors' tendency to tilt their portfolios toward local stocks. First, positive local economic shocks could simultaneously cause high local portfolio returns and increase entry. If returns on local stocks (which tend to form a large part of the local stock portfolios) are affected by changes in the prospects of the local economy, local portfolio returns might be associated with increases in local wealth. The increase in wealth might induce some of the non-participants to enter the stock market. To address this alternative explanation, we test for the underlying assumption that changes in regional stock market wealth are related to changes in non-participants' wealth. Our estimates do not suggest a significant relation, which makes it difficult to reconcile the results with local wealth shocks. In addition, the neighborhood return effect is nonlinear, as discussed above, which has no natural explanation under the local wealth shocks hypothesis.

According to the second alternative explanation, non-participants with limited attention might more closely follow local stocks rather than the overall stock market, and decide to buy local stocks after observing good performance. We test this hypothesis by analyzing stock market entry through Initial Public Offerings (IPO). Prior to the listing, these companies are not part of any local portfolio, and it is not possible to follow their performance. We find the neighborhood return effect also in stock market entries through IPOs, which is not consistent with the alternative explanation. Finally, we also find the neighborhood effect in areas that have no local companies with listed stocks. This refutes any alternative explanation involving a special role for local companies.

The contribution of the paper is summarized as follows. First, we establish the importance of outcome-based social learning in a large scale statistical study. Second, we find that communication is selective, only good outcomes affect behavior. Third, we find evidence broadly consistent with extrapolative expectations. In addition to these general results, the paper also contributes to the understanding of the dynamics of stock market participation.[5] The dynamic

[5] Stock market participation has attracted attention from both academics and policy makers (see, e.g., Mankiw and Zeldes, 1991; Haliassos and Bertaut, 1995). Most of the earlier evidence on stock market participation comes from cross-sectional snapshot data (an exception is Brunnermeier and Nagel (2008) who study dynamics of individual

aspect of participation and investor communication is emphasized by Cao, Coval, and Hirshleifer (2002). In their model, agents facing fixed costs of trading may delay trading until price movements validate their private signals. In a yet more narrow sense, we also contribute to the determinants of investor demand in Initial Public Offerings (IPO). This is a byproduct from an additional analysis on IPO data which we use as a robustness check.

In the remainder of the paper, Section 2 reviews the literature on different forms of social interaction and develops the hypotheses. Section 3 discusses relevant institutional background and the data sources, and Section 4 presents the empirical strategy. Section 5 presents the results, while section 6 assesses alternative explanations. Section 7 concludes.

2. Literature and hypotheses

2.1. Earlier literature on social interaction

Empirical studies on social interaction come from various fields and settings. One manifestation of the magnitude of the studies is the vocabulary used in the studies: social mechanism can be referred to as peer effects, community effects, neighborhood effects, network effects, herding, mimicking, conformity, or observational learning. A common problem in empirical studies is identification: a finding that individuals' choices are related to peers' choices is not necessarily due to social interaction (Manski, 1993 and 2000). Many empirical studies nevertheless argue, and use varying identification strategies to show, that social interaction is indeed driving the observed relation between average behavior of a peer group and an individual's behavior.

Evidence for action-based social learning has been found in farmers' crop choices (Foster and Rosenzweig, 1995), criminal activity (Glaeser, Sacerdote, and Scheinkman, 1996), labor market participation of married women (Woittiez and Kapteyn, 1998), use of welfare benefits (Bertrand, Luttmer, and Mullainathan, 2000), membership of social groups (Sacerdote, 2001), participation in a pension plan (Duflo and Saez, 2002), participation in the stock market (Hong, Kubik, and Stein, 2004), choice of a health plan (Sorensen, 2006), automobile purchases

---

wealth and participation). Cross-sectional studies of social interaction find a static positive relation between the participation decisions by individuals and their communities (Hong, Kubik, and Stein, 2004; Brown et al., 2008).

(Grinblatt, Keloharju, and Ikäheimo, 2008), choice of workplace (Topa, Bayer, and Ross, 2009), and choice of dishes from a restaurant menu (Cai, Chen, and Fang, 2009).

Sacerdote (2001) and Zimmerman (2003) show that academic performance in college is positively affected by the performance of a randomly assigned roommate. One interpretation of this result is that having been assigned to a high-achieving roommate helps in emulating good study practices. This is consistent with outcome-based social learning. It may, however, be due to other externalities, such as motivation or competitive pressure.

2.2. Outcome-based social learning

Early work by social psychologists has shown the importance of observing outcomes as children copy new behaviors (Bandura and Walters, 1963). In the field of animal studies, Call and Tomasello (1994) find that orangutans do not merely copy what other orangutans are doing when they are trying to obtain out-of-reach food. Instead, they adopt techniques that yield best results, paying attention to both orangutan and human demonstrators' success. Theorizing based on these findings informs the development of hierarchical models of learning in the behavioral brain sciences (for a review see Byrne and Russon, 1998). As mentioned earlier, economic theorists have also modeled outcome-based social learning (Ellison and Fudenberg 1993 and 1995; McFadden and Train 1996; Persons and Warther 1997, and Banerjee and Fudenberg 2004).

Despite the theoretical interest, empirical research on the outcome-based dimension of social learning has been limited. The studies that we are aware of are in the areas of agricultural and development economics. Munshi (2004) finds that Indian farmers planted more of a new high yielding variant of wheat in the early 1970s if farmers in the neighboring districts had received good yields from that variant. His farm-level results are based on a single snapshot of data in 53 villages. Kremer and Miguel (2007) test for peer effects in the decision to undergo drug treatment against intestinal worms in Kenya. There can be large benefits to society via less infections, if enough people take the treatment. Despite this positive externality, Kremer and Miguel find that people are less likely to take the deworming drugs if their peers have taken these drugs. This suggests that people are learning something about the outcomes of their peers, although Kremer and Miguel are not able to measure those outcomes. Conley and Udry (2009) investigate fertilizer use in 47 pineapple farms in the Akwapim South district of Ghana at a time when pineapple was a new produce in the area. Figuring out the right amount of fertilizer requires some experimentation, as the optimum amount depends on local conditions. Conley and Udry's results show that the farmers' use of fertilization is affected by the amounts used, as well as the profits achieved by their peers.

2.3. Extrapolative expectations and selective communication

Our main contribution is to examine the dynamics of social interaction: what circumstances make observational learning more likely to take place and to be more effective. We hypothesize that outcomes experienced by peers are a key driver on both dimensions. That is, social interaction is more likely to take place when peers' outcomes have been good and the signal from observing the outcome is more effective in altering individual behavior after good outcomes.

In addition to our main hypothesis about the role of outcome-based learning, we investigate the channels through which outcomes affect social learning. For this purpose, it is useful to sketch two simplified hypotheses. According to the first hypothesis, people form their expectations about stock returns based on their peer's experiences. Under this hypothesis, people have full information about whether their peer is participating in the stock market, but do not know about the investment outcomes. Knowledge about outcomes becomes available when people talk, and it exerts an incremental influence in forming expectations beyond other sources of information, such as the media. This hypothesis is called the 'extrapolative expectations' hypothesis.

The second hypothesis relaxes the assumption of full information about the peers' stock market participation decisions. Under this hypothesis, an individual is not aware of the peers' stock market participation decision, but becomes aware of it when people talk. People do not use peer outcomes to update expectations. However, talking makes people become aware of the stock market participation of their peers. People may want to imitate their peers due to a 'Keeping up with the Joneses' effect (Abel, 1990; Gali, 1994; Bakshi and Chen, 1996). Such an effect may result from preference-based interactions, such as conformity to social norms (Akerlof, 1976), or competition for resources (DeMarzo, Kaniel, and Kremer, 2004). This hypothesis is called the 'selective communication' hypothesis.

There are several reasons to expect that people are more likely to talk about their stock market experiences after they have experienced good returns. First, people may simply enjoy discussing their positive stock market experiences more than the bad ones. Second, there can be social benefits to appearing to be a competent investor. This provides an incentive to keep quiet about one's poorly performing investments. Third, various theories in psychology (under the conceptual umbrellas of motivated cognition, self-deception, or attribution) predict, and experiments confirm, that people have a self-serving bias in recalling and interpreting the factors involved in their successes and failures. People tend to emphasize their personal role when they achieve good performance, while emphasizing external factors with poor performance. Cognitive dissonance theory (Festinger, 1957; see Akerlof and Dickens, 1982, for an economic model) argues that a discrepancy between one's actions and self-image causes discomfort, and that people try to act and think in ways that reduce the discomfort. Bénabou and Tirole (2002) present a general economic model in which agents protect their self-esteem by engaging in self-deception through selective memory and awareness.

Both 'extrapolative expectations' and 'selective communication' hypotheses can alone account for a positive relation between past neighborhood returns and the tendency of new investors to enter the market. The selective communication hypothesis makes a further testable prediction that is not shared by the 'extrapolative expectations' hypothesis: the relation between past peer outcomes and future actions should be stronger with good outcomes compared to bad outcomes. The hypotheses are not mutually exclusive, however. In particular, it is possible that communication is selective, but once communication takes place, naïve extrapolation still occurs.

3. Institutional background and data

3.1. Stock market participation in Finland

Direct stock ownership has been the primary means for stock market participation in Finland during our sample period. From 1995 to 2002, the stock market participation rate has increased from 9.3% to 13.9%, which implies an annual increase of 50 basis points. The largest increases occurred in 1998 to 2000, and they coincide with high market returns and many equity offerings that attracted new investors to participate in the stock market. Privatizations of government-owned companies played an important role in increasing stock market participation—individuals made about 240,000 subscriptions in these offerings (Keloharju, Knüpfer, and Torstila, 2008).

Other means to participate in the stock market comprise the government-sponsored obligatory pension plan, voluntary pension products, and mutual funds. All employees in Finland are automatically included in a government-sponsored defined benefits pension plan. There are no personal pension accounts such as 401k, and an individual employee has no influence on the amount of her own contribution, nor the selection of investments in these government-sponsored plans. Pensions are mostly financed by the contributions of the current workforce.

It is also possible to make additional voluntary pension investments with some tax benefits. Numbers from the beginning of 2003 suggest that 15% of the assets in these voluntary plans have been allocated to financial products in which the return depends on the performance of the capital markets.[6] For the remaining 85% of the assets the return is linked to money market rates.

Mutual funds are a relatively recent phenomenon in Finland. The first mutual funds were introduced in 1987, but their use by households remained limited for several years. In the beginning of our sample period in 1995, households' assets held in all types of mutual funds were 11% of the households' direct stock market investments. This figure increased to 32% by 2002, the end of our sample period. Discussions with bankers reveal that it was not very common for people without any direct stock holdings to purchase equity mutual funds during the 1990s. Since 2002 the popularity of mutual funds has grown.

We lack accurate statistics on the amount invested in equity mutual funds, but their share has been about 30% to 50% of all mutual fund assets during 1998-2002. We assume that the same applies to voluntary pension plans. Based on these assumptions, one can approximate the relative importance of various channels for stock market investments. Of all the forms of households' equity exposure (directly held stock, mutual funds, pension plans), directly held stock has been the dominant form: its average share during the sample period has been 89% to 93%.

3.2. Data

The data comes from an established source for investor level data. The data set is derived from the Finnish Central Securities Depository (FSCD), an official registry that includes every stock market transaction of every stock market participant in the whole Finnish stock market. The data span a time period from January 1995 to November 2002. The data also include a number of

[6] The estimates in this section are based on data provided by the Finnish Bankers' Association and the Finnish Association for Mutual Funds.

investor characteristics. For our purposes, the most important is the place of residence, which is available at three points in time: January 10, 1997, June 30, 2000, and November 27, 2002.

From this data set, we extract two types of data:

a. Entry dates. For every investor in the sample, we determine the stock market entry date as the first day on which an investor buys stocks of publicly listed companies. We require that there are no other transactions with positive volume on that day to exclude entries through equity offerings, gifts, inheritances, divorce settlements, and others that do not represent a genuine active stock market entry decision. This definition of stock market entry leaves also out investors that have a stock market position in the beginning of the sample period. It is possible that some of the investors that enter the stock market during our sample period may have already participated in the stock market earlier, but have exited the market before the beginning of our sample period.

b. Neighborhood returns. In the absence of a direct mapping from an individual to his or her neighbors, we use zip codes as the neighborhoods. In total, there are about 2,700 zip codes in Finland at the end of 2002. For each zip code in the sample, we define neighborhood return as the value weighted average return on the portfolio that the investors residing in a zip code held in the beginning of a month. We also use the equally weighted portfolio return in some of the analysis.

We merge this dataset with socioeconomic census data from Statistics Finland. Table 1, Panel A, summarizes descriptive statistics of the zip codes in the sample. The average stock market entry rate implies that on average 1.6% of the inhabitants of a zip code enter the stock market during the sample period. For simple illustrative purposes, Panel B reports results of regressions of the determinants of stock market entry rates aggregated at the zip code level. Entry rates are higher in areas with higher wealth and income, and higher levels of education. Members of the Swedish-speaking minority, which hold a disproportionately high proportion of wealth, are also more likely to enter the stock market. Figure 1 plots the stock market entry rates across the whole country. Stock market entry rates are much higher in Southern and Western Finland, reflecting the concentration of urban areas in the South and the Swedish-speaking communities in the West.

4. Methods and identification

We begin the analysis by introducing a major improvement compared to earlier studies on social interaction. We show that assigning a positive experience with stocks to randomly selected individuals increases the number of new shareholders in their neighborhood in the following months compared to control areas where no assignment took place. Such a difference-in-difference estimation is made possible by a demutualization of a large telecom firm that turned tens of thousands individuals from non-participants to shareholders. It is necessary to note, however, that this analysis is unable to address our main hypothesis directly, because there is no variation in the outcomes experienced by the randomly assigned individuals.

In an effort to test our main hypothesis, we turn to an analysis of changes in zip code entry rates at monthly level. This brings in significant cross-sectional and temporal variation in our key variables of interest, neighborhood returns and entry rates. Our identification of the effect of experienced outcomes on neighborhood entry relies on the dynamic feature of the data. Another important feature is our ability to unambiguously identify the shareholders and non-shareholders in an area at any point in time. Since we lack register data on individuals who never became shareholders during the sample period, we aggregate the number of shareholders to zip codes, for which we have the total number of inhabitants available. This aggregation helps us to define the dependent variable in most of the regressions as the number of new investors per the total number of inhabitants, i.e., the change in the stock market participation rate in a zip code. The change variable enables us to trace the changes in the exact number of participating and non-participating individuals at any point in time.

We construct a panel of observations consisting of 2,668 cross-sections (zip code areas) and 93 periods (calendar months). This is not a typical panel data set in that it has an unusually large number of both cross-sections and time periods. In estimating the models, we rely on methods developed in political science for analysis of political connections between countries over time. Beck and Katz (1995; 2004) show that such models can be estimated with simple OLS techniques and more complex techniques provide minor or non-existing improvements.

The main regression model is the following:

$$y_{it} = a + b(\mathbf{w}_{i,t-1}\mathbf{r}_{i,t-1}) + y_{i,t-1} + p_{i,t-1} + u_{it} \tag{1}$$

where the subscripts i and t refer to zip code areas and months, respectively. The dependent variable, y, is the (log) number of new investors entering the stock market in the zip code during a month. **w** is a row vector of local portfolio weights in each stock in the beginning of a month, **r** is a column vector of returns for those stocks during a month, so that **wr** is the return of the local (zip code) portfolio. p is the stock market participation rate, and u is the error term that absorbs the fixed effects for months and zip codes used in the analysis. Standard errors are calculated allowing clustering at the zip code level.

The outcome variable **wr** is our main variable of interest. Lagged number of new investors and lagged stock market participation are included to control for trends in the popularity of stock market entry. According to our main hypothesis neighborhood return measured as **wr** should be positively related to the number of entries.

Our empirical model is designed to capture pure social interaction in which communication takes place between two groups of people living in the same area (shareholders and non-participants). This structure, combined with the significant amount of temporal and cross-sectional variation, rules out alternative mechanisms based on reverse causality and common unobservables. Consider first reverse causality, the possibility that stock market entry in a particular area causes higher neighborhood returns, not vice versa. There could be price pressure from new stock market participants to stocks owned by existing investors in an area[7]. While this mechanism might affect the contemporaneous relation between entry and return, it is not possible that future actions would affect past prices. This alternative mechanism is therefore ruled out by only using the lagged neighborhood return as a regressor.

The second objection to our methodology is the problem with common unobservables: there might be mechanisms that cause the local portfolio to perform consistently better than the market and simultaneously encourage more investors to participate in the stock market. One example of such an influence could be varying levels of financial sophistication. In areas with higher level of

[7] Such price pressure is rather unlikely, however, given that the average number of new investors entering the market in a month in a zip code is 0.19.

financial sophistication existing investors may do well in their stock market investments and new investors are encouraged to participate in the stock market not because of good returns, rather because of good quality advice from their peers.[8] Such effects are eliminated in our analysis, as the influence of common time-invariant unobservables is swamped by the zip code fixed effects.

The third objection to our methodology is based on local time-varying shocks and limited attention towards local stocks. The impact of local wealth shocks may arise from unexpected increases in profitability of local companies combined with the fact that individual investors own disproportionately local stocks. This is a reasonable candidate for an explanation for a positive relation between past returns and stock market entry. If the profitability of local listed companies were to spill over to the local economy and increase the wealth of residents of an area, this might make some of the non-participants switch to a positive allocation to stocks, either due to decreased risk aversion or lower proportional cost of participation (Abel, 2001; Vissing-Jørgensen, 2003). We also consider a potential story where investors follow only local companies and decide to participate after good returns on local stocks. However, the evidence presented in Section 6 goes against these alternative explanations.

In this section we have argued that our empirical framework is immune to many of the econometric challenges that plague empirical studies of social interaction. In influential papers, Manski (1993, 2000) has argued that it is impossible to identify social interaction in most of the available data sets, due to reverse causality, common unobservables, and common responses to shocks. We believe that our data and method is an exception in the literature as it relates the outcomes of one group to the actions of another, distinct group. It also has the dynamic feature which enables addressing the common unobservables critique.

[8] Studies on individual investors' investment performance suggest that outperformance is very rare within individual investors. Most individuals would be better off by investing in a passive market portfolio as they underperform the market by trading excessively and making bad stock picks (Odean 1999; Barber and Odean 2000).

5. Results

5.1. Natural experiment

We start by analyzing a natural experiment that investigates how the behavior of prospective stock market entrants changed as a result of a random assignment of stock to some investors. This analysis is an improvement over previous studies of social interaction, where random assignment has not been available. Studies without such a feature may suffer from the reverse causality and common unobservables problems identified by Manski (1993, 2000).

The experiment involves a telecom company by the name of Elisa, which changed its organizational form from a mutual customer owned company into a corporation owned by shareholders. In the old organizational form, a mutual share (customer certificate) entitled the customer to a fixed telephone line. Alternatively, prospective customers were able to rent the certificate. In the 1990s the company developed new businesses, such as rapidly growing wireless services, and the mutual corporate form was no longer seen as appropriate. As the companies demutualized and went public, the customer certificates were converted into listed shares. This conversion created tens of thousands of new stock market participants. The shares of Elisa were listed on July 1, 1999. The owner of one old customer certificate received 150 new shares, valued at about 3,000 euros at the end of the first day of trading. The price of the customer certificate obtained from the mutual company had for some time been 925 euros.

The stock had been trading on a "pre-list" gray market earlier, but not during two months before the listing. It is likely that the listing on the main list refocused the owners attention to the stock and its valuation, providing topics for discussion. Some of them may have only become aware of the stock's value after the listing, which would provide a happy surprise.

This positive experience is directly relevant for our main hypothesis: we expect to see more stock market entry in areas were people received the shares compared to control areas. The increase in the number of new shareholders could be driven by two factors. First, the increase through the conversion can have a socially transmitted externality in making stock market participation more prevalent in the treated areas. Second, the increase can be due to the good stock market experiences of the newly minted shareholders, which should increase communication and facilitate the formation of positive expectations. In this experiment, we are unable to distinguish between these two mechanisms, but we will return to this issue later.

To measure the effect of the Elisa conversion we use a difference-in-differences (DID) design proceeding as follows. We measure the changes in entry rates (the number of new investors / number of non-investors) around the conversion in both treatment and control areas, where the treatment areas are the ones with new stock market entries through the conversion.

To establish a benchmark we measure entry rates in every zip code during the three months before the listing date. The three-month numbers are converted to monthly figures by dividing by three, as they will be compared to monthly entry rates after the listing date. We use two measures in calculating the number of new investors: those who enter the market by any active decision (but leaving out passive entries such as bequests), and, a narrower version, those who enter the market by a regular purchase of a listed stock. In calculating these measures we exclude investors who later will receive the treatment shares from both the numerator (new investors) and the denominator (the number of non-investors). In essence, we investigate changes in entry behavior by people who are not investors, and who will not become investors by receiving the treatment shares.

We then measure entry rates for each of the three months following the listing of the converted shares, i.e., July, August, and September of 1999. By definition, each new stock market entry comes now from someone who was not an investor before, and did not receive the treatment. In the affected areas we thus analyze the behavior of non-participants who did not own a customer certificate. In the areas not treated we measure entry rates of all non-investors.

One may wonder whether the treated groups are different from the control group. Because the definition for treatment is based on owning a customer certificate, it turns out that the non-owners are likely to be younger, less wealthy, and have lower income. We address this selection problem in the regressions by controlling for age, wealth, and income at the zip level. With imperfect controls there is likely to remain some systematic differences between the treated and non-treated groups. To the extent that the non-treated group has a lower intrinsic propensity to enter the stock market (which is likely to be the case), we conclude that being not able to fully control for characteristics of the treatment and control group biases our estimates downwards.

The difference-in-differences estimator is constructed as follows. The data is a panel with four observations for each zip code (the three months prior to the event aggregated together, as well as each of the following three months entered separately). 'Treatment' –dummy takes the value of one for treated zip codes, and zero otherwise. 'Month 1', 'Month 2', and 'Month 3' take the value of one in their designated months after the event, and zero otherwise. The interaction terms with the Treatment dummy and the three month dummies produce the difference-in-differences estimator broken down by each month.

Table 2 shows the base case results. The indicator variable for the treated zip areas, i.e., ‘Treatment’, is significantly negative (or not significantly positive) in all specifications. The entry rates in these areas are therefore generally lower. Note that in these areas we only consider the behavior of people who did not own the customer certificate, as the certificate owners by definition automatically become shareholders. This result is therefore consistent with the idea that people who did not own the certificate are less likely to participate in the market due to their lower wealth, as discussed earlier.

The DID estimators, i.e., the interactions between the ‘Treatment’ dummy and the monthly dummies after the event, that measure the increased entry due to the treatment, are positive and significant in all cases. The first column shows the results for all active forms of participation, including equity issues. The DID estimator enters the regression significantly stronger using this definition compared to restricting the sample to regular stock market buys. This is generally true for other specifications and sample variations as well. This result is consistent with IPOs being an important gateway for new investors to enter the market.

In the following we concentrate on the results obtained from regular buys only, as the marketing of equity issues may have a geographical bias that is not easy to control for. The next four columns in Table 2 show the results considering investors entering the stock market through ‘regular stock market buys’. In the full model in Column 2 the DID estimator for Month 1 has a coefficient of 2.45×10^{-4}, implying that the first month increase in the stock market entry rate is 2.5 basis points. Given that zip code level entry rates are about 50 basis points per year, the effect corresponds to a 60% increase in the monthly entry rate. The spillover effect from entry through the conversion declines as one moves in time further from the event. For example, the DID estimator for the third month corresponds to a 1.1 basis point increase in the entry rate.

Our results are insensitive to the inclusion of control variables. In Column 3 we include no control variables. In Column 4 we add income and wealth and in Column 5 age and education. In all the specifications, the first month DID estimator equals 2.5. basis points with other months entering with very similar coefficients to the base case model. We also consider subsamples. First, we restrict the analysis to densely populated urban areas only, namely areas with at least

700 inhabitants per square kilometer. This is a very harsh restriction, leaving out almost 90% of the observations. The effect remains significant for the first two months after the event, but the magnitude is about 2/3 of that in the base case. Second, we limit to areas with at least one new investor in the three-month period prior to the event. The magnitude is now about twofold to the base case effect. Third, we use an alternative way to scale the variables: instead of percentages, the variables are specified as log (1 + number of new investors). Log (1 + population) is also added as a control variable. The effect is significant for the first two months, but not for the third.

These results provide strong support for the idea that exogenous increases in stock market participation have a causal positive spillover effect on non-participants: higher stock market participation of geographically close peers increases individual's propensity to enter the stock market. Because our results come from an experiment where the peer experience with the stock market was positive and identical, we know turn to separating the pure spillover effect from the social learning hypothesis.

5.2. Past neighborhood returns and stock market entry

We separate the outcome-based learning from action-based learning by analyzing a panel of stock market entry decisions. Specifically, in each month and each zip code, we explain the number of new investors entering the stock market with the returns experienced by existing investors. This regression is discussed in detail in Section 3.1.

Table 3, Panel A reports descriptive statistics of the key variables in the regression. On average 0.19 investors enter the stock market in a zip code in a month. The average return on the portfolio of existing investors equals 1.2% and the mean stock market participation rate is 9.6%. There is considerable variation in these numbers both across time and across zip codes. For example, half of the monthly zip code returns fall between -3.9% and 5.7%. Stock market entry is a rare event: in most cases there are no new investors entering the market, although the maximum number of new investors equals 42. Stock market participation rates are also low with half of the observations falling between 5.4% and 11.3%.

Table 3, Panel B, reports the results of the regression (1). Column 1 leaves out the lagged number of new investors and lagged participation rate while Column 2 adds the lagged number of investors. The full model appears in Column 3.

All specifications provide strong support for our main hypothesis, i.e., that past peer outcomes affect individual actions. In our full model past neighborhood return enters the regression with a significantly positive coefficient of 8.5×10^{-2}. Lagged variables suggest that there is strong autocorrelation in stock market entry and that stock market entries are positively related to the level of stock market participation in a zip code. Thus leaving out the lagged control variables changes somewhat, but not dramatically, the coefficient on the neighborhood return variable.

We assess economic significance by providing marginal effects of the coefficients and comparing the effect of neighborhood return to the effect of market returns. Marginal effects are calculated as a change in the number of investors resulting from a one-standard deviation change in the neighborhood return. Our benchmark in this analysis is the average zip code, which has 225 investors and a 1.2% monthly return. Based on our estimates the typical increase in the logged number of investors in a month equals $0.085 \times 0.012 = 0.001$. This translates into 0.2 new investors. With one-standard deviation increase in the return from its mean, the increase in the log number of investors equals $0.085 \times 0.097 = 0.016$, which corresponds to 1.9 new investors. This suggests that the effect of increasing the return by one-standard deviation has a eightfold effect on stock market entry.

How sensible are these estimates of the effect of neighborhood returns? One way to assess the validity of our coefficient estimates is to compare them to changes in the aggregate stock market participation rate. In an average month, the stock market participation rate has increased in the US from 1989 to 2001 by 0.13 percentage points (Curcuru et al., 2009). Our estimates imply that the increase in the stock market participation rate due to social learning in a typical month with a 1.2% return translates into an increase of 0.2 new investors. The average zip code with 225 investors and a stock market participation rate of 9.6% will thus experience a 0.08 percentage point monthly increase in the stock market participation rate due to social learning. This estimate is reasonable when compared with the monthly change in the US aggregate stock market participation rate.

Column 4 in Table 3, Panel C, drops the month fixed effects and replaces them by a variable measuring market return in the previous month. Although we lose the ability to control for time effects, this analysis allows us to compare the strength of the neighborhood return effect to that of the market returns. However, we need to note that market returns are imperfect controls for time effects, as the month fixed effects probably also to some extent pick up performance differences between stocks typically owned by individuals and the market index. For example, suppose the weight of stock *A* is higher in all the neighborhood portfolios compared to the market index. This is possible as the neighborhood portfolio weights are constructed from individuals' holdings, leaving out institutions. Superior performance by stock *A* would then lead to a systematic positive effect on all neighborhood portfolios compared to the market index, and new stock market entries would be positively affected throughout the country. Fortunately, this mechanism does not interfere with our comparison of the neighborhood returns to market returns.

The results are consistent with the idea that market returns are imperfect controls of time effects. For example, the coefficients of the lagged dependent variable and participation rate increase significantly, both in magnitude and statistical significance, suggesting that part of the imperfectly controlled time effects spill over to these variables.

The results show that neighborhood return effects are more important than market returns: the coefficient on neighborhood return is about 30% higher than that of market returns. Market returns are highly visible in the news media as well as in the publications of stockbrokers and mutual fund companies. The neighborhood returns, on the other hand, are not published anywhere. Instead, they are transmitted by word-of-mouth. This is consistent with the idea that vivid stories of actual experiences by one's peers have a greater effect on behavior compared to general statistical information.

5.3. Extrapolative expectations and selective communication

In Table 4 we break the neighborhood return into positive and negative parts. The results show that the effect on entry rates comes exclusively from positive returns. The coefficient on negative returns is actually negative in columns 1 and 2, meaning that negative returns increase entry as well. However, the effect is statistically insignificant and economically small. Column 3 adds the participation rate variable. This produces coefficient for negative returns that is essentially zero.

Table 4, Panel B, shows a graph of the effect of neighborhood return broken down to categories representing five percentage point intervals between –10% and +15%, as well as categories for a return less than –10% and in excess of +15%. We choose the cut points for the extreme categories so that approximately 5% of the distribution falls within them both in the right and left tail of the distribution.

The effect of returns is zero for negative returns, significantly positive for positive returns, and particularly strong for returns in excess of 15%. This pattern is consistent with the selective communication hypothesis outlined in Section 2. That is, people are more likely to discuss their stock market experiences with others when those experiences have been favorable. Importantly, the pattern is not explained by the naïve extrapolation hypothesis alone. Extrapolative expectations without selective communication would lead to a symmetric relation: people would get couraged by the good experiences of their peers, but they would also get discouraged by bad experiences, and thus poor returns would reduce market entry by new investors. However, these hypotheses are not mutually exclusive. It is possible that communication is selective, but naïve extrapolation could still occur when the information is transmitted.

There can be social benefits to appearing to be a competent investor. One could appear competent by discussing successful investment experiences. This would provide a motivation to misrepresent one's investment success. However, as we have shown, actual neighborhood returns predict new entry, so the results are not explained by the idea that people simply make up successful investment stories. Thus, people can be sincere about their performance given their information set, but discussing good experiences may have a higher payoff compared to discussing poor ones. Self-serving bias in recall and attention can also cause people to talk about their investments more when the performance is good. Bénabou and Tirole (2002) show that this kind of behavior can be a result of (boundedly) rational optimization.

This kind of selective communication about financial performance may potentially explain other financial market phenomena as well. For example, Chevalier and Ellison (1997) and Sirri and Tufano (1998) find a convex flow-performance relation for mutual funds: good returns increase net flows substantially, but poorly performing funds do not suffer a comparable decrease in flows. The asymmetry in fund flows can be explained by fund companies increasing marketing efforts in response to good historical returns (Jain and Wu, 2000) and by fund companies discarding personnel or techniques that produce underperformance (Lynch and Musto, 2003). Our results suggest that, if word-of-mouth effects are important in mutual fund choice, then a greater tendency to share information about one's successful mutual fund purchases may play a part in producing a convex flow-performance relation.

5.4. Performance of new investors' trades

We investigate the performance of the new stock market entrants to learn more about the nature of the information that the new investors acquire. Existing studies find that individual investors make value-reducing trading decisions, while institutional investors gain at their expense (Odean, 1999; Barber and Odean, 2000; Barber et al., 2009). To measure the performance for a particular investor, we only consider the stocks bought on the entry date. In other words, we do not track possible future changes in the portfolio in this analysis. We calculate the value weighted returns for this entry portfolio for holding periods of 20, 60, 125, and 250 trading days.

If those investors who enter the market after their neighbors have earned good returns are prone to naïve extrapolation, then they might be more susceptible to other behavioral biases, and thus more likely to make inferior stock picking choices as well. Under this hypothesis, the new investors' returns should be negatively related to the past neighborhood return. We run regressions in which the dependent variable is this entry portfolio return, and the explanatory variable is the neighborhood return over the past month. We include month and stock fixed effects.

Table 5 shows the results. The performance of the new investors' initial portfolio is strongly negatively related to the past neighborhood return. Most of the negative effect comes in the course of the first month (20 trading days). The underperformance grows up to the 6-month holding period, and then reverses slightly for the 1-year period. Results for longer holding periods (not reported) show no signs of further reversal. If anything, the performance tends to deteriorate. A one standard deviation (11%) increase in the neighbors' past return causes approximately a two percent annual underperformance for the new investors. These results imply that investors who enter the market after their neighbors enjoyed good returns, tend to buy precisely those stocks which underperform from that point forward. It shows that investors are not learning tactically useful information by observing their neighbors' performance. The negative relation is consistent with naïve extrapolation.

It is still possible that the net impact of the neighbors' success on the new investors' welfare is positive. The short-term losses must be weighed against the long-term gains of stock market participation. The new investors could later on learn to diversify better, thereby eliminating any further underperformance. The initial underperformance could also turn out to be discouraging, however. Investors are known to overweight personally experienced investment outcomes (Kaustia and Knüpfer, 2008), so the disappointing initial experiences could drive the new investors out of the market. These ideas provide interesting questions for future research, but they are beyond the scope of this paper.

6. Alternative explanations

6.1. Local wealth shocks

There are alternative explanations that may also predict a positive relation between local portfolio returns and the number of new investors. A positive relation between past neighborhood returns and future market entry may arise from local shocks to real economy if there are spillovers from local stock market wealth to other components of local wealth. Consider, for example, a case in which the prospects of local firms improve unexpectedly and the market value of these companies, mostly held by local investors, increases. This would show up in the local portfolio as a positive return. At the same time the positive shock to the local economy might affect non-participants' wealth in the form of higher demand for local products and services, or higher salaries and bonuses paid to employees of local companies. There might also be wealth effects from the stock market to stockholders' consumption (e.g., Poterba, 2000), which might also generate spillovers to local economy.

The increase in local non-stock market wealth might induce some of the non-participants to switch to shareholders, either due to changes in risk aversion, or lower per-period costs of participation (Abel, 2001; Vissing-Jørgensen, 2003). Brunnermeier and Nagel (2008) show that agents with either constant relative risk aversion (CRRA) and fixed per-period participation costs or habit formation preferences should be more likely to enter the stock market when their wealth increases. They find a positive relation between changes in wealth and stock market entry in microdata.

Some ex ante skepticism may be warranted about the ability of local wealth shocks to generate the results of this paper. This is because it should take more time than just one month for a shock in the real economy to cause changes in local non-stock market wealth. We show that the neighborhood effect works at a relatively short horizon. In unreported analysis, we find that lagged returns beyond one month are not statistically significantly related to stock market entry.[9] Thus, to explain the results, the impact of the economic shocks would have to propagate very rapidly in the real economy and have only transitory effects.

We nevertheless also directly test the underlying assumption of the local wealth shocks hypothesis, i.e., that changes in local stock market wealth are correlated with changes in non-participants wealth. We use municipality level wealth data from 1994 to 2002, obtained from Statistics Finland.[10] For each municipality and each year, we divide total wealth into two components: stock market wealth calculated from the investor data, and a residual non-stock market wealth. We adjust the wealth measures by subtracting net stock market inflows from stock market wealth, and adding them back to the residual wealth. This adjustment effectively removes the fluctuations in wealth components that are due to rebalancing between different asset classes. We then divide the wealth measures by the number of inhabitants with taxable income to control for year-to-year shifts in population. We calculate the growth rates of both stock market wealth and residual wealth by dividing the flow-adjusted end-of-year wealth measure by beginning-of-year raw wealth measure. This gives growth rates in the absence of flows between asset classes and population changes.

The local wealth shock story assumes that the growth of stock market wealth is positively correlated with the growth of non-stock market wealth. To investigate this assumption, we regress the growth of adjusted residual non-stock wealth (Δn_{it}) on the growth of adjusted stock market wealth (Δs_{it}). We include municipality and year fixed effects to account for unobservable time-invariant municipality characteristics and wealth shocks that concurrently affect the whole economy. Standard errors are adjusted for clustering at the municipality level (443 municipalities). Some small municipalities have very small values of wealth, resulting in extreme

[9] We also consider daily frequency and conclude that neighborhood returns less than month before the entry decision have a considerable positive effect on stock market entry. These results come from an investor-level analysis where we randomly assign one day from t=-370 to t=-120 before the actual entry date as a shadow-entry date. The actual entry date and the shadow-entry date are coded as the dependent variable taking the value of one and zero, respectively. The independent variables are the neighborhood returns measured over intervals -1 to -30, -31- to -60, -61 to -90, and -91 to -120 days.

[10] These data are available only at the level of municipality and annual frequency. Every municipality includes at least one zip code.

growth rates. We exclude these observations by discarding observations that fall outside the 1st and 99th percentiles of the growth rate distribution.

We obtain the following coefficient estimates with the *t*-values in parentheses:

$$\underset{}{\Delta n_{it}} = \underset{(4.28)}{0.181} + \underset{(0.22)}{0.007}\Delta s_{it} \quad N = 3{,}195 \quad R^2 = 0.006 \tag{2}$$

According to the estimates, the growth of local residual wealth is not related to the growth of local stock market wealth. The sign of the coefficient is positive, but it is statistically, as well as economically, insignificant. These results are inconsistent with the key assumption of the local wealth shocks hypothesis.

Further evidence against local wealth shocks comes from our results on the nonlinearity of the neighborhood effect, which has a natural explanation in social learning and communication. The local wealth shocks hypothesis, in contrast, does not explain why only positive returns would be related to stock market entry. Taken together, we believe that local wealth shocks are not driving the result between neighborhood returns and stock market entry.

6.2. Local stocks and home bias

Investors tend to overweight local stocks in their portfolios (Coval and Moskowitz, 1999; Huberman, 2001; Grinblatt and Keloharju, 2001). This tendency may lead prospective investors to more closely follow local stocks, either due to information advantage, a preference for familiarity, or simply limited attention. If prospective investors more closely follow local stocks, they may end up buying local stocks after seeing those stocks produce high returns. Given that ownership in local stocks generates some of the observed cross-sectional variation in neighborhood returns, cases in which local stocks have been performing well are likely to coincide with a high return on the corresponding neighborhood portfolio. New investors would thus enter the stock market after observing good performance in the local stocks, irrespective of observing any outcomes of their neighbor's portfolio.

We investigate this alternative explanation by analyzing market entry in stocks with no prior history, namely Initial Public Offerings (IPO). In essence, these stocks do not exist before the listing. They are not part of the local portfolio, and it is not possible to follow their performance before the listing. Therefore, according to the limited attention hypothesis, past local returns should not influence market entry via IPO stocks.

In Table 6 we perform an analysis of the demand for IPO stocks. The sample is a time-zip code panel with the full set of zip codes that was used in the main analysis of stock market entry. Time is now defined as the start date of the subscription period of an IPO. Details of the sample IPOs and their characteristics appear in Kaustia and Knüpfer (2008).

The dependent variable in the regression is the log number of investors participating in a particular IPO in a zip code. Independent variables are similar to those in Table 3, and they are measured at the beginning of the subscription period of each IPO. The neighborhood return is measured from a period of 30 days before the beginning of the subscription period. The regressions are estimated by including zip code fixed effects. We control for time effects with IPO fixed effects, which pick up the influence of market returns and other contemporaneous effects. The results show that the likelihood of entering the stock market via an IPO is strongly affected by past neighborhood returns. The fact that the results on IPO subscriptions are similar to the main results cannot be explained by the tendency to follow local stocks.

Further evidence against any special role of local firms in local portfolios comes from an analysis that repeats the baseline results in Table 3 by excluding municipalities that have no local companies with listed stocks. The analysis includes 197,511 observations and yields a coefficient estimate of 6.27×10^{-2} (*t*-value 5.42). The estimates are very similar to the baseline specification, which is strong evidence against local stocks driving our results.

Besides providing a possibility to address an important alternative explanation, the results on IPOs make an independent contribution to the IPO literature. Derrien (2005) and Ljungqvist, Nanda, and Singh (2006) model the impact of investor sentiment on IPO demand and pricing patterns. Empirical studies find that these demand and pricing patterns are correlated with measures of investor sentiment (Lee, Shleifer, and Thaler, 1991; Rajan and Servaes, 1997; Lowry, 2003; and Cornelli, Goldreich, and Ljungqvist, 2006). Kaustia and Knüpfer (2008) argue that investors' past personal experiences with IPOs provide a microfoundation for the role of sentiment in IPO demand.

The results of our paper show how the IPO demand by inexperienced investors is influenced by positive stock market experiences of their peers. Good performance by retail investors thus introduces a positive externality in the form of new investors, something that is valued by the issuers and their investment banks (Cook, Kieschnick, and Van Ness, 2006). On the other hand, there may be negative consequences as well, as the excess demand by inexperienced investors can contribute to the formation of asset price bubbles.[11]

7. Conclusion

This paper documents a tendency for a specific type of social learning in the stock market: the returns experienced by the existing investors in a neighborhood in a given month encourage new investors to enter the stock market in the following month. This type of social learning has been called 'adaptive emulation' in sociology, and also 'goal emulation' in animal studies. We present what to our knowledge is the first large scale statistical evidence on the issue. Our data covers all individual investors in Finland, and we control for neighborhood and time fixed effects. A natural experiment also shows that people tend to enter the stock market after their neighbors became stockholders as the local mutually held phone company converted its customer certificate into exchange listed shares.

A common criticism toward studies on social interaction is that the effects may be caused by similar characteristics of the people in a community (contextual interactions and correlated effects in the taxonomy of Manski, 2000). Two features of this setting ensure that this criticism does not apply. First, we study two distinct and well-defined groups of people, existing investors and new investors. Second, we are able to measure the outcomes, not only the actions, of the existing investors.

The neighborhood return effect is asymmetric, and only positive returns are related to entry. A plausible explanation is that people do not communicate when the returns are negative. People may enjoy discussing their good stock market experiences more, and would rather not discuss the bad ones. This is consistent with the model of Bénabou and Tirole (2002) in which agents protect their self-esteem by engaging in self-deception through selective memory and awareness.

[11] See Mitchell, Pulvino, and Stafford (2002) and Lamont and Thaler (2003) for evidence of blatant violations of the law of one price in the case of parent company and subsidiary stocks. See Smith, Suchanek, and Williams (1988) and Porter and Smith (1994) for the role of investor experience in asset pricing bubbles in laboratory experiments. Greenwood and Nagel (2009) find that less experienced mutual fund managers showed trend-chasing behavior in their technology stock holdings around the technology stock bubble of the late 1990s.

We also find that the new investors' performance is negatively related to their neighbors' past return. Taken together, the results are consistent with social epidemics of financial type feeding on success stories that spread in a network of peers. Naïve extrapolation from these stories can lead to bad decisions, as well as play a part in explaining phenomena such as localized real estate booms and busts, Ponzi type securities scams, and the dismal stock market timing record of individual investors.

There are good reasons to expect that our results apply to other areas of life. Learning about others' outcomes often requires communication. For example, consider the purchase of a new product. Selective communication is likely to apply to new product innovations, especially when individuals feel responsibility over their decisions and want to protect their self-esteem. Outcome-based learning may even be stronger in settings other than the stock market, where unobserved peer characteristics and a large stochastic component in the outcome are key features. These features are strong impediments to social learning, so outcome-based social learning may have a more powerful influence in other, less uncertain environments.

References

Abel, Andrew A., 1990, "Asset Prices under Habit Formation and Catching up with the Joneses", *American Economic Review* 80, 38-42.

Abel, Andrew A., 2001, "The Effects of Investing Social Security Funds in the Stock Market When Fixed Costs Prevent Some Households from Holding Stocks", *American Economic Review* 91, 128-148.

Akerlof, George A., 1976, "A Theory of Social Custom, of Which Unemployment May Be One Consequence", *Quarterly Journal of Economics* 94, 749-775.

Akerlof, George A. and William T. Dickens, 1982, "The Economic Consequences of Cognitive Dissonance", *American Economic Review* 72, 307-319.

Bakshi, Gurdip and Zhiwu Chen, 1996, "The Spirit of Capitalism and Stock Market Prices", *American Economic Review* 86, 133-157.

Bandura, Albert and Richard H. Walters, 1963, "Social Learning and Personality Development", New York: Holt, Rinehart, & Winston.

Banerjee, Abhijit and Drew Fudenberg, 2004, "Word-of-Mouth Learning", *Games and Economic Behavior* 46, 1-22.

Barber, Brad M., Yi-Tsung Lee, Yu-Jane Liu, and Terrance Odean, 2009, "Just How Much Do Individual Investors Lose by Trading?", *Review of Financial Studies* 22, 609-632.

Barber, Brad M. and Terrance Odean, 2000, "Trading Is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors", *Journal of Finance* 55, 773-806.

Beck, Nathaniel and Jonathan Katz, 1995, "What to Do (and Not to Do) with Time-Series Cross-Section Data", *American Political Science Review* 89, 634–647.

Beck, Nathaniel and Jonathan Katz, 2004, "Time Series Cross Section Issues: Dynamics", Working Paper.

Bénabou, Roland and Jean Tirole, 2002, "Self-Confidence and Personal Motivation", *Quarterly Journal of Economics* 117, 871-915.

Bertrand, Marianne, Erzo F.p. Luttmer, and Sendhil Mullainathan, 2000, "Network Effects and Welfare Cultures", *Quarterly Journal of Economics* 115, 1019-1055.

Brown, Jeffrey R., Zoran Ivković, Paul A. Smith, and Scott J. Weisbenner, 2008, "Neighbors Matter: Causal Community Effects and Stock Market Participation", *Journal of Finance* 63, 1509-1531.

Brunnermeier, Markus, and Stefan Nagel, 2008. "Do Wealth Fluctuations Generate Time-Varying Risk Aversion? Micro-Evidence on Individuals' Asset Allocation", *American Economic Review* 98, 713-736.

Byrne, Richard W. and Anne E. Russon, 1998, "Learning by Imitation: A Hierarchical Approach", *Behavioral and Brain Sciences* 21, 667-721.

Cai, Hongbin, Yuyu Chen, and Hanming Fang, 2009, "Observational Learning: Evidence from a Randomized Natural Field Experiment", *American Economic Review*, Forthcoming.

Call, Josep and Michael Tomasello, 1994, "The Social Learning of Tool Use by Orangutans (Pongo Pygmaeus)", *Human Evolution* 9, 297-313.

Cao, H. Henry, Joshua D. Coval, and David Hirshleifer, 2002, "Sidelined Investors, Trading-Generated News, and Security Returns", *Review of Financial Studies* 15, 615-648.

Chevalier, Judith and Glenn Ellison, 1997, "Risk Taking by Mutual Funds as a Response to Incentives", *Journal of Political Economy* 105, 1167-1200.

Coleman, James S., Elihu Katz, and Herbert Menzel, 1966, "Medical Innovation: A Diffusion Study", New York: Bobbs-Merrill.

Conley, Timothy G. and Christopher R. Udry, 2009, "Learning About a New Technology: Pineapple in Ghana", American Economic Review, Forthcoming.

Cook, Douglas O., Robert Kieschnick, and Robert A. Van Ness, 2006, "On the Marketing of IPOs", *Journal of Financial Economics* 82, 35-61.

Cornelli, Francesca, David Goldreich, and Alexander Ljungqvist, 2006, "Investor Sentiment and Pre-IPO Markets", *Journal of Finance* 61, 1187-1216.

Coval, Joshua D. and Tobias J. Moskowitz, 1999, "Home Bias At Home: Local Equity Preference in Domestic Portfolios", *Journal of Finance* 54, 2045-2074.

Curcuru, Stephanie, John Heaton, Deborah Lucas, and Damien Moore, 2009, "Heterogeneity and Portfolio Choice: Theory and Evidence", In Y. Ait-Sahalia and L. P. Hansen (eds.), Handbook of Financial Econometrics, Elsevier.

Demarzo, Peter M., Ron Kaniel, and Ilan Kremer, 2004, "Diversification as a Public Good: Community Effects in Portfolio Choice", *Journal of Finance* 59, 1677-1715.

Derrien, François, 2005, "IPO Pricing in "Hot" Market Conditions: Who Leaves Money on the Table?", *Journal of Finance* 60, 487-521.

Duflo, Esther and Emmanuel Saez, 2002, "Participation and Investment Decisions in a Retirement Plan: The Influence of Colleagues' Choices", *Journal of Public Economics* 85, 121-148.

Ellison, Glenn and Drew Fudenberg, 1993, "Rules of Thumb for Social Learning", *Journal of Political Economy* 101, 612-643.

Ellison, Glenn and Drew Fudenberg, 1995, "Word-of-Mouth Communication and Social Learning", *Quarterly Journal of Economics* 110, 93-125.

Festinger, Leon, 1957, "A Theory of Cognitive Dissonance", Stanford University Press.

Foster, Andrew D. and Mark R. Rosenzweig, 1995, "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture", *Journal of Political Economy* 103, 1176-1209.

Gali, Jordi, 1994, "Keeping Up with the Joneses: Consumption Externalities, Portfolio Choice, and Asset Prices", *Journal of Money, Credit and Banking* 26, 1-8.

Glaeser, Edward L., Bruce Sacerdote, and José A. Scheinkman, 1996, "Crime and Social Interactions", *Quarterly Journal of Economics* 111, 507-548.

Greenwood, Robin and Stefan Nagel, 2009, "Inexperienced Investors and Bubbles", *Journal of Financial Economics*, Forthcoming.

Griliches, Zvi, 1957, "Hybrid Corn: An Exploration in the Economics of Technological Change", *Econometrica* 25, 501-522.

Grinblatt, Mark and Matti Keloharju, 2000, "The Investment Behavior and Performance of Various Investor Types: A Study of Finland's Unique Data Set", *Journal of Financial Economics* 55, 43-67.

Grinblatt, Mark and Matti Keloharju, 2001, "How Distance, Language and Culture Influence Stockholdings and Trades", *Journal of Finance* 56, 1053-1073.

Grinblatt, Mark, Matti Keloharju, and Seppo Ikäheimo, 2008, "Social Influence and Consumption: Evidence from the Automobile Purchases of Neighbors", *Review of Economics and Statistics* 90, 735-753.

Haliassos, Michael and Carol Bertaut, 1995, "Why Do So Few Hold Stocks?", *Economic Journal* 105, 1110–1129.

Hong, Harrison, Jeffrey D. Kubik, and Jeremy C. Stein, 2004, "Social Interaction and Stock-Market Participation", *Journal of Finance* 59, 137-163.

Huberman, Gur, 2001, "Familiarity Breeds Investment", *Review of Financial Studies* 14, 659-680.

Jain, Prem C. and Joanna Shuang Wu, 2000, "Truth in Mutual Fund Advertising: Evidence on Future Performance and Fund Flows", *Journal of Finance* 55, 937-958.

Kaustia, Markku and Sami Torstila, 2008, "Political Preferences and Stock Market Participation", Working Paper, Helsinki School of Economics.

Kaustia, Markku and Samuli Knüpfer, 2008, "Do Investors Overweight Personal Experience? Evidence from IPO Subscriptions", *Journal of Finance* 63, 2679-2702.

Keloharju, Matti, Samuli Knüpfer, and Sami Torstila, 2008, "Do Retail Incentives Work in Privatizations?", *Review of Financial Studies* 21, 2061-2095.

Kremer, Michael and Edward Miguel, 2007, "The Illusion of Sustainability", *Quarterly Journal of Economics* 122, 1007-1065.

Lamont, Owen A. and Richard H. Thaler, 2003, "Can the Market Add and Subtract? Mispricing in Tech Stock Carve-Outs", *Journal of Political Economy* 111, 227-268.

Lee, Charles M. C., Andrei Shleifer, and Richard H. Thaler, 1991, "Investor Sentiment and the Closed-End Puzzle", *Journal of Finance* 46, 75-109.

Ljungqvist, Alexander, Vikram Nanda, and Rajdeep Singh, 2006, "Hot Markets, Investor Sentiment, and IPO Pricing", *Journal of Business* 79, 1667-1702.

Lowry, Michelle, 2003, "Why Does IPO Volume Fluctuate So Much?", *Journal of Financial Economics* 67, 3-40.

Mankiw, N. Gregory and Stephen P. Zeldes, 1991, "The Consumption of Stockholders and Nonstockholders", *Journal of Financial Economics* 29, 97-112.

Manski, Charles F., 1993, "Identification of Endogenous Social Effects: The Reflection Problem", *Review of Economic Studies* 60, 531-542.

Manski, Charles F., 2000, "Economic Analysis of Social Interactions", *Journal of Economic Perspectives* 14, 115-136.

Mitchell, Mark, Todd Pulvino, and Erik Stafford, 2002, "Limited Arbitrage in Equity Markets", *Journal of Finance* 57, 551-584.

Munshi, Kaivan, 2004, "Social Learning in a Heterogenous Population: Technology Diffusion in the Indian Green Revolution", *Journal of Development Economics* 73, 185-213.

Odean, Terrance, 1999, "Do Investors Trade Too Much?", *American Economic Review* 89, 1279-1298.

Persons, John C. and Vincent A. Warther, 1997, "Boom and Bust Patterns in the Adoption of Financial Innovations", *Review of Financial Studies* 10, 939-967.

Poterba, James, 2000. "Stock Market Wealth and Consumption", *Journal of Economic Perspectives* 14, 99-118.

Porter, David P. and Vernon L. Smith, 1994, "Stock Market Bubbles in the Laboratory", *Applied Mathematical Finance* 1, 111-127.

Rajan, Raghuram G. and Henri Servaes, 1997, "Analyst Following of Initial Public Offerings", *Journal of Finance*, 52, No. 2, 507-529.

Sacerdote, Bruce, 2001, "Peer Effects with Random Assignment: Results for Dartmouth Roommates", *Quarterly Journal of Economics* 116, 681-704.

Shiller, Robert J., 1984, "Stock Prices and Social Dynamics", *Brookings Papers on Economic Activity* 1984:2, 457-498.

Shiller, Robert J., 1990, "Market Volatility and Investor Behavior", *American Economic Review* 80, 58-62.

Sirri, Erik and Peter Tufano, 1998, "Costly Search and Mutual Fund Flows", *Journal of Finance* 53, 1589-1622.

Smith, Vernon L., Gerry L. Suchanek, and Arlington W. Williams, 1988, "Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets", *Econometrica* 56, 1119-1151.

Sorensen, Alan T., 2006, "Social Learning and Health Plan Choice", *RAND Journal of Economics* 37, 929-945.

Topa, Giorgio, Patrick Bayer, and Stephen L. Ross, 2009, "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market", *Journal of Political Economy*, Forthcoming.

Vissing-Jørgensen, Annette, 2003, "Perspectives on Behavioral Finance: Does 'Irrationality' Disappear with Wealth? Evidence from Expectations and Actions", *NBER Macroeconomics Annual*.

Woittiez, Isolde and Arie Kapteyn, 1998, "Social Interactions and Habit Formation in a Model of Female Labour Supply", *Journal of Public Economics* 70, 185-205.

Zimmerman, David J., 2003, "Peer Effects in Academic Outcomes: Evidence from a Natural Experiment", *Review of Economics and Statistics* 85, 9-23.

Table 1

Descriptive statistics

This table reports descriptive statistics of zip code level socioeconomic characteristics and stock market entry rates. Panel A summarizes the socioeconomic characteristics of the 2,649 zip codes at the end of the sample period (in 2002). 'Stock market entry rate' is the proportion of inhabitants entering the stock market during the sample period 1995 to 2002. 'Population density' is the number of inhabitants divided by the area of a zip code (in km^2). Age measures the average age of all the inhabitants of a zip code. 'College degrees' is the proportion of people with higher academic education. 'Swedish-speaking' measures the proportion of people whose mother tongue is Swedish (Finland has two official languages, Finnish and Swedish). Income and wealth are based on official tax filings of all the individuals living in a zip code. Panel B explains stock market entry rates with socioeconomic characteristics. The regressions include decile dummies (except one) for population density, not reported for brevity. Heteroskedasticity robust *t*-values are reported in parentheses below coefficients. The table multiplies the original regression coefficients by 100.

Panel A: Socioeconomic characteristics of zip codes							
Variable	Mean	Sd	Minimum	25%	Median	75%	Maximum
Stock market entry rate (%)	1.6	1.2	0.0	1.0	1.5	2.0	35.0
Number of inhabitants	1,911	2,801	101	272	609	2,423	24,734
Area (km^2)	110	226	0.1	21	55	114	3,511
Population density (per km^2)	299	1,010	0.1	4	10	59	23,464
College degrees (%)	55.4	10.0	25.0	48.0	55.0	62.0	92.0
Swedish-speaking (%)	7.0	21.3	0.0	0.0	0.0	1.0	98.0
Income (1000€)	15.6	4.1	9.1	13.0	15.0	17.3	66.4
Wealth (1000€)	52.6	25.2	10.0	39.0	48.0	60.0	397.0

Panel B: Regressions of stock market entry rate			
Dependent variable	Stock market entry rate		
Specification	1	2	3
Ln (Income)	0.41	1.11	
	(1.85)	(4.72)	
Ln (Wealth)	0.50		0.58
	(6.14)		(6.78)
College degrees	0.01	0.01	0.02
	(2.80)	(1.67)	(4.38)
Swedish-speaking	0.01	0.01	0.01
	(7.66)	(7.59)	(8.14)
Number of observations	2,649	2,649	2,649
Adjusted R^2	0.15	0.14	0.15

Table 2

Natural experiment

The results of a natural experiment on how an individual's likelihood to enter the stock market is influenced by an exogenous shock to the number of stock market participants in the neighborhood. The dependent variable is (Number of new investors / Number of non-stockholders) in a zip code during a month. The experiment is based on a conversion of customer certificates to common stock in a local telecom demutualization on July 28, 1999. Entry rates (the number of new investors / number of non-investors) in zip code areas are measured three months before and after the listing date. In Column 1, the number of new investors is calculated as the number of investors who enter the market by any active decision (leaving out e.g. gifts and bequests). In Columns 2 through 5, the number of new investors is based on those individuals who enter the market by a regular purchase of some listed stock. 'Treatment' is a dummy variable set to one for zip codes in which the conversions took place, zero otherwise. The interaction terms between the treatment dummy and the three month dummies produce the difference-in-differences estimator broken for each month. Control variables are income, wealth, age, and education at the zip code level. 'Income' is the median income of individuals, 'Wealth' is the mean taxable wealth of families. 'Age' refers to six continuous variables that measure the proportion of inhabitants falling into each age category from 18-24, 25-34, 35-44, 45-54, 55-64, and 65- years. 'College degrees' is the proportion of people with higher academic education. The *t*-values (in parentheses below coefficients) are robust at the zip code level. The regressions are estimated with OLS. The table reports the original regression coefficients multiplied by 10,000.

Specification	All entries	Stock market buys			
	1	2	3	4	5
Treatment × Month 1	6.70	2.45	2.45	2.45	2.45
	(5.76)	(3.39)	(3.29)	(3.35)	(3.38)
Treatment × Month 2	4.63	1.57	1.57	1.57	1.57
	(6.65)	(3.63)	(3.45)	(3.57)	(3.63)
Treatment × Month 3	3.86	1.05	1.05	1.05	1.05
	(6.44)	(2.69)	(2.55)	(2.65)	(2.68)
Treatment	-2.64	-1.49	0.30	-1.14	-1.31
	(-5.55)	(-4.54)	(1.14)	(-3.62)	(-4.19)
Month 1	0.23	-0.28	-0.28	-0.28	-0.28
	(0.78)	(-1.22)	(-1.21)	(-1.22)	(-1.22)
Month 2	-1.06	-0.94	-0.94	-0.94	-0.94
	(-4.28)	(-4.55)	(-4.52)	(-4.55)	(-4.56)
Month 3	-1.14	-1.08	-1.08	-1.08	-1.08
	(-4.88)	(-5.82)	(-5.77)	(-5.82)	(-5.83)
Control variables					
Income and wealth	Yes	Yes	No	Yes	Yes
Age	Yes	Yes	No	No	Yes
College degrees	Yes	Yes	No	No	No
Adjusted R^2	0.09	0.03	0.01	0.03	0.03
Number of observations	10,580	10,580	10,580	10,580	10,580

Table 3

Past neighborhood returns and stock market entry

The results of regressions of the number of investors entering the stock market in a month in a zip code. The dependent variable is Ln (1 + Number of new investors). 'Neighborhood return' is defined as the return on the portfolio of all investors in a zip code, calculated as the sum of the return on stocks held in a zip code weighted by the value of holdings in each stock in the beginning of a month. 'Participation rate' is the number of stock market participants divided by the number of inhabitants in a zip code. 'Market return' is the return on the Helsinki Exchanges Portfolio Index that represents the whole Finnish stock market. The regressions 1 to 3 include fixed effects for months and zip codes, column 4 drops the month fixed effects and adds market return as an explanatory variable. The regressions are estimated with OLS and *t*-values (in parentheses below coefficients) are robust at the zip code level. The table reports the original regression coefficients multiplied by 100. All regressions have 251,823 observations and include zip code (2,649 zip codes) fixed effects.

Panel A: Descriptive statistics

	Mean	Sd	Minimum	25%	Median	75%	Maximum
Ln (1 + Number of new investors)	0.18	0.42	0.00	0.00	0.00	0.00	3.76
Neighborhood return (%)	1.17	8.56	-47.67	-3.88	0.96	5.72	99.92
Participation rate (%)	9.55	6.85	0.00	5.42	7.94	11.30	100.00

Panel B: Regressions

Dependent variable	Ln (1 + Number of new investors)			
Specification	1 (OLS)	2 (LDV)	3 (LDV)	4 (LDV)
Neighborhood return	9.00	8.10	8.51	6.09
	(7.15)	(6.77)	(7.01)	(8.42)
Lagged dependent variable		22.05	21.91	31.57
		(39.11)	(38.63)	(52.55)
Participation rate			24.63	103.21
			(4.01)	(22.44)
Market return				4.72
				(7.23)
Month fixed effects	Yes	Yes	Yes	No
Zip code fixed effects	Yes	Yes	Yes	Yes
Number of zip codes	2,649	2,649	2,649	2,649
Number of observations	251,823	251,823	251,823	251,823
Overall R^2	0.11	0.29	0.30	0.31

Table 4

The relation between past neighborhood returns and stock market entry for positive and negative returns

This table presents the results of a regression that is similar to Table 2, except that the neighborhood return accounts for an asymmetric relation between returns and entry. The dependent variable is Ln (1 + Number of new investors). In Panel A, the neighborhood return is replaced by a piecewise linear functional form employing a single change in the slope at the return equaling zero. In Panel B, the return is divided into 8 categories, with the omitted category in the regression being the return between 0% and 5%. All the coefficients are multiplied by 100.

Panel A: Regressions

Dependent variable	Ln (1 + Number of new investors)		
Specification	1 (OLS)	2 (LDV)	3 (LDV)
Max (Neighborhood return, 0)	15.32	13.52	13.14
	(7.15)	(6.90)	(6.71)
Min (Neighborhood return, 0)	-2.37	-1.63	0.18
	(-0.85)	(-0.65)	(0.07)
Lagged dependent variable		22.05	21.91
		(39.11)	(38.63)
Participation rate			24.04
			(3.92)
Month fixed effects	Yes	Yes	Yes
Zip code fixed effects	Yes	Yes	Yes
Number of zip codes	2,649	2,649	2,649
Number of observations	251,823	251,823	251,823
Overall R^2	0.11	0.29	0.30

Panel B: Return broken down into categories

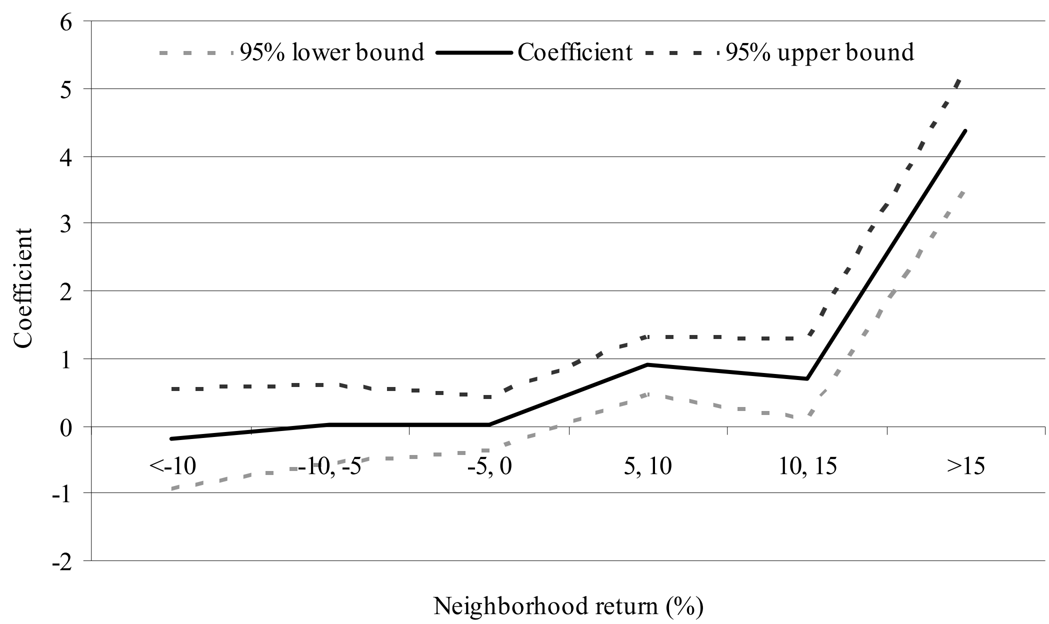

Table 5

The performance of the new investors' initial portfolios

The results of regressions of the new investors' return on past neighborhood return. That data is investor level, and the sample consists of investors who enter the market using a regular stock exchange purchase transaction. Portfolios are formed as of the entry date, value weighting each stock that is purchased during that day. The percentage capital gains of the portfolios are calculated for the holding periods indicated in the table column headings. 'Neighborhood return' is defined as the monthly return on the portfolio of all investors in a zip code, calculated as the sum of the return on stocks held in a zip code weighted by the value of holdings in each stock, one month prior to the entry date of each investor. The regressions are estimated with OLS and include fixed effects for each calendar month and for each stock.

Dependent variable	New investor return			
Specification	Holding period in trading days			
	20	60	125	250
Neighborhood return	-0.17	-0.22	-0.25	-0.20
	(18.00)	(13.26)	(13.85)	(7.51)
Month fixed effects	Yes	Yes	Yes	Yes
Stock fixed effects	Yes	Yes	Yes	Yes
Number of observations	93,860	93,066	91,285	86,692
Overall R^2	0.29	0.49	0.57	0.55

Table 6

Past neighborhood returns and stock market entry through IPOs

This table reports the results of the regressions on the number of investors entering the stock market through an IPO. The dependent variable is Ln (1 + Number of new investors). There are 57 IPOs in the sample. Independent variables are from the beginning of the subscription period of an IPO. 'Neighborhood return' is defined as the return on the portfolio of all investors in a zip code, calculated as the sum of the return on stocks held in a zip code weighted by the value of holdings in each stock in the beginning of a month. 'Participation rate' is the number of stock market participants divided by the number of inhabitants in a zip code. The regressions are estimated with OLS, with Columns 2 and 3 including the lagged dependent variable as an explanatory variable. The regressions include fixed effects for IPOs and zip codes, and *t*-values (in parentheses below coefficients) are robust at the zip code level. The table reports the original regression coefficients multiplied by 100. All regressions have 152,076 observations.

Panel A: Descriptive statistics							
	Mean	Sd	Minimum	25%	Median	75%	Maximum
Ln (1 + Number of new investors)	0.11	0.43	0.00	0.00	0.00	0.00	5.24
Neighborhood return (%)	2.18	10.58	-58.46	-3.98	1.97	8.85	164.48
Participation rate (%)	9.60	6.69	0.25	5.43	7.98	11.41	51.20

Panel B: Regressions			
Dependent variable	Ln (1 + Number of new investors)		
Specification	1 (OLS)	2 (LDV)	3 (LDV)
Neighborhood return	9.14	8.16	8.28
	(7.80)	(7.06)	(7.13)
Lagged dependent variable		3.86	3.78
		(10.28)	(9.97)
Participation rate			20.97
			(6.86)
Number of zip codes	2,649	2,649	2,649
Number of observations	150,993	150,993	150,993
Overall R^2	0.28	0.29	0.30

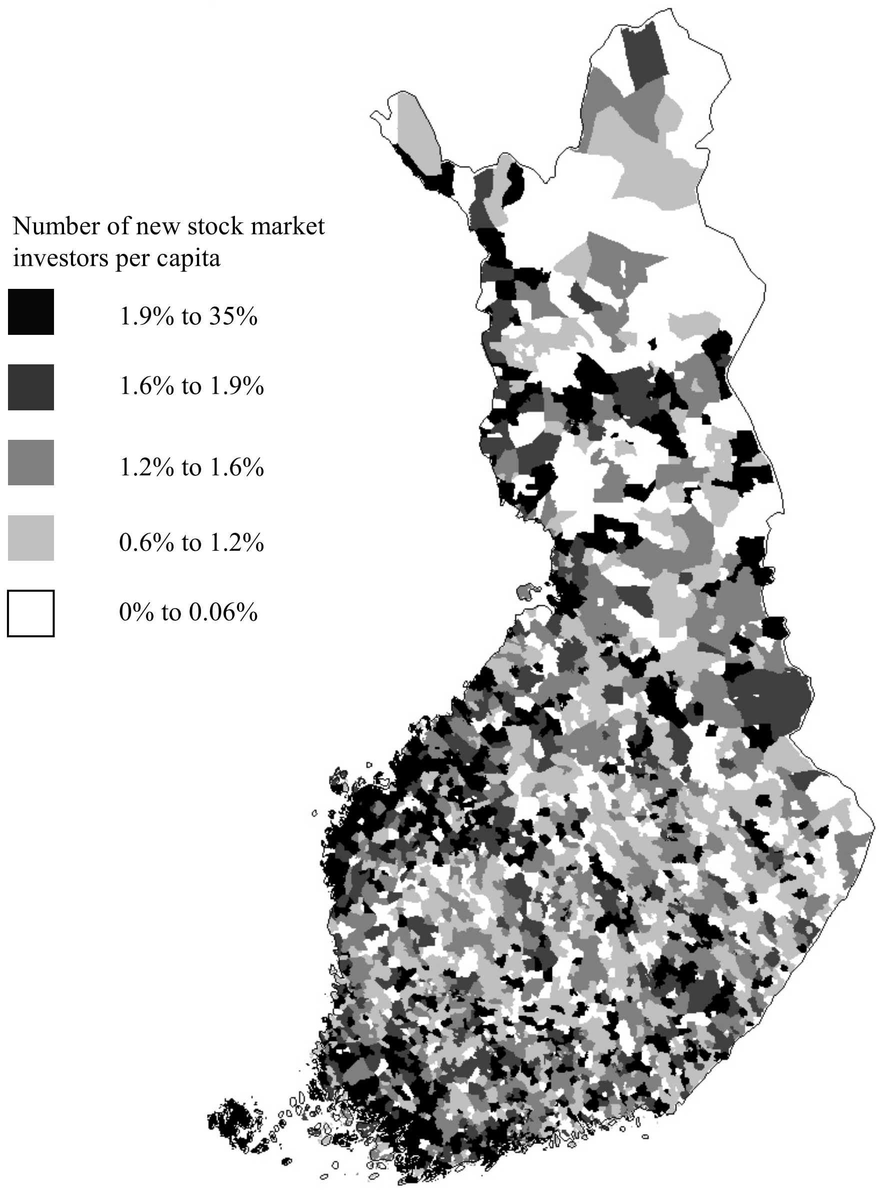

Figure 1. Distribution of stock market entry rates across zip codes in Finland. This figure plots the stock market entry rates, i.e., the number of new investors entering the stock market during the sample period 1995 to 2002 divided by the total number of inhabitants at the end of 2000, across Finnish zip codes.